Army Research Laboratory

Interval Scales From Paired Comparisons

by Andrew A. Thompson

ARL-TR-6003 May 2012

NOTICES

Disclaimers

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5069

ARL-TR-6003 **May 2012**

Interval Scales From Paired Comparisons

Andrew A. Thompson
Weapons and Materials Research Directorate, ARL

REPORT DOCUMENTATION PAGE						*Form Approved* *OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**						
1. REPORT DATE *(DD-MM-YYYY)* May 2012		**2. REPORT TYPE** Final				**3. DATES COVERED (From - To)** 1 January 2011–1 October 2011
4. TITLE AND SUBTITLE Interval Scales From Paired Comparisons						**5a. CONTRACT NUMBER**
						5b. GRANT NUMBER
						5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Andrew A. Thompson						**5d. PROJECT NUMBER**
						5e. TASK NUMBER
						5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-WML-A Aberdeen Proving Ground, MD 21005-5069						**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TR-6003
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)						**10. SPONSOR/MONITOR'S ACRONYM(S)**
						11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT The development of scales can increase the objectivity in a domain characterized by subjectivity. The increase in objectivity typically leads to improved decisions and enhances the effectiveness of an organization or an investigation. The issues associated with the use of paired comparisons to form interval scales are the focus of this report.						
15. SUBJECT TERMS paired comparison, interval scale, Bradley-Terry model						
16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**		**19a. NAME OF RESPONSIBLE PERSON** Andrew A. Thompson
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UU	22		**19b. TELEPHONE NUMBER** (*Include area code*) (410) 278-6805

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Acknowledgments

The author would like to thank Dave Webb for the time he took to review this report. His insightful comments and attention to detail improved the quality and readability of this report.

INTENTIONALLY LEFT BLANK.

1. Introduction

The use of scales within the scientific community is widespread. Many of these scales are based on comparisons of some property. For example, the Mohs hardness scale is based on the ability of one sample of matter to scratch another. Scales are objective when they are based on an observable property; however, sometimes the scales are based on human evaluation, and there is conflict in the results. Subjective scales can be used to evaluate product preferences or summarize expert opinion. For example, scales can be developed to quantify benefits of competing avenues of research or the impact of different technologies related to a specific goal. The development of a technology impact scale can bring objectivity and quantification to a subjective domain. Ordinal or interval scales can be developed to summarize expert opinion, material properties, or utility. As artificial intelligence (AI) applications become more prevalent, Turing tests can be based on scale development. By creating a pool of AI and human responses, an interval scale can be made by asking individuals which responses are more human. An analysis of the scale values would determine if a system passes the Turing test.

The development of scales can increase the objectivity in a domain characterized by subjectivity. The increase in objectivity typically leads to improved decisions and enhances the effectiveness of an organization or an investigation. The issues associated with the use of paired comparisons to form interval scales are the focus of this report.

2. Background

Paired comparisons offer a direct way to present items for evaluation according to a specific criterion. The term " item" refers to the objects being compared and can therefore refer to a multitude of possible instantiations. The method has been in use for over 150 years and was used extensively by the psychophysicists in the 1800s; these experiments involved the human perception of physical stimuli (e.g., light intensity, sonic pitch, sound intensity, taste, smell, etc.). Many of these experiments attempted to define the just noticeable difference (JND) for human perception of the stimuli. Defining the JND for sensory items was a central concern of psychophysics. Each comparison can be thought of as a contest. By asking the subject to focus on one comparison in isolation, paired comparison typically produces reproducible data of high quality. Since each observation gives an ordinal relationship, a consistent response would not violate any transitive orderings. The term "items" will refer to the objects being compared. By noting the number of transitive relations that are inconsistent, a measure of evaluator or subject consistency can be formed. A drawback to the method is the number of comparisons that need to be made. For n objects n(n-1)/2 comparisons are required, the number of comparisons for each

evaluator is of the order n squared. For example, if 15 descriptions are compared on a criterion, there are 105 pair-wise comparisons for each evaluator. If it is further supposed that there are 480 items and each has had 15 descriptions generated, then 480*105 paired comparisons are required. The data requirements can be demanding; however, this is offset by the quality of the data obtained. In addition to quantifying subjective data, paired comparisons can also be used to summarize contest data where one side wins (e.g., a rod defeating an armored plate, one sports team defeating another team, or the preference of one product or feature over another). The summarization of paired comparisons is typically presented as ordinal rankings of the scaled items. While an ordinal ranking of the items is often sufficient for the task, some problems require interval scales. Under specific assumptions, interval scales can be developed for the items. Typically, the scale values are assumed to have the same dispersion. Paired comparisons can be used to create scales for everything from the subjective to the objective.

Generalized Linear Models (GLMs) provide the background material for interval scale development, i.e., scale development can be conceived as an application of GLM. John Nelder and Robert Wedderburn (1972) formulated GLMs as an extension of ordinary linear models to include error distributions that could be put in a specific exponential form. This increased the number of problems that could be treated using the developed theory associated with ordinary linear models. To use the theory, the underlying distribution must be able to be transformed to fit a specific exponential form. The explanatory variables are related to the parameters of the underlying distribution through a link function. Link functions can be interpreted as a change of variables or as a change of scale. The explanatory variables are estimated by minimizing the deviance of the particular model.

For paired comparisons, the data distribution is binomial (win or loss), and the link function determines the interpretation of the scale. Typically, a logistic or probit function provides a viable link function. The logistic function can be used to convert the logarithm of odds into a probability. The probit function is the inverse of the standard normal cumulative distribution function. For a given probability, the probit function will return the z-value associated with that probability. A source of information of GLMs containing many examples has been prepared by Dobson and Barnett (2008). The estimated distance between two items will be based on the probability that one exceeds the other; if the paired comparisons result in a probability of 0.5, the items are considered similar and will have the same scale value. For a probit or logistic link, probabilities of 0 and 1 give estimates of negative and positive infinity (off the scale).

In figure 1, colors are used to represent distinct items or people. The distribution of scores could indicate the skill with which a certain team or individual will perform. In another sense, it may represent the perceived value of an item, with the variation representing the lack of precision of the evaluator or perhaps the lack of agreement between evaluators. Note that while on the average, the blue (second peak from left) will be inferior to the cyan (third peak from left); there will be a significant number of times where the blue is preferred over the cyan. A zone of mixed results between blue and cyan will typically be established after a small number of trials.

Contrast this to the situation between the green (first peak) and red (fourth peak); in this case, it will take many trials for a single green win (or preference) to occur (a result not leading to an infinite scale value). While it can be difficult to establish the distance between items far apart, it does not take as many samples to determine the distance between items that are close together. This observation can serve as the basis for reducing the number of paired comparisons to be made. In these cases, the inconsistent data or zones of mixed results are the basis of proximity determination. For a completely consistent set of data, an ordinal ranking is the only reasonable summary; interval scales cannot be determined. Many papers state that the paired comparisons need to be uncorrelated. However, Mosteller (1951) shows that the uncorrelation requirement is not always necessary.

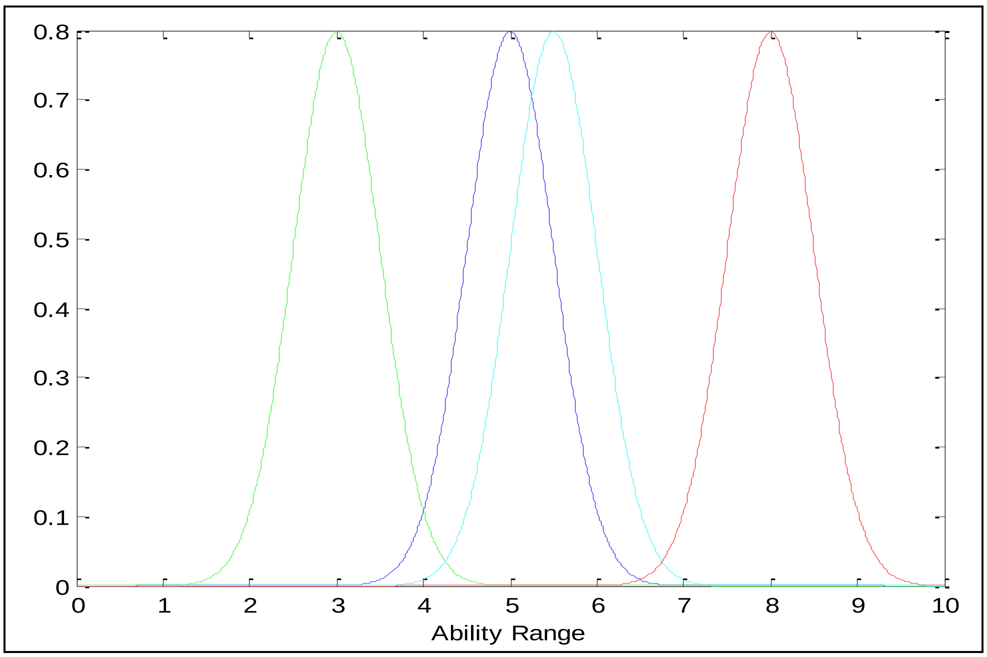

Figure 1. Ability distributions for four items.

3. Bradley-Terry Models

Although Bradley-Terry models were in existence 20 years before GLMs were developed, they can be interpreted as an application of GLMs to paired comparisons. This domain of problems is large enough to warrant its own subfield, and specialized software with specific data formats has been developed. The typical Bradley-Terry model involves binomial data with a logistic link function. In some cases, probit link functions are used. There are packages for Bradley-Terry models in R (Firth, 2005). When using a logistic link, the difference in the assigned scores give the odds of preference or victory. For a probit link, the difference in the scores is interpreted as the z-score associated with the probability of winning. In either case, the actual values are not

important, but the differences are crucial. These scales are interval scales because there is not an absolute zero; transposition can be done for convenience. Agresti (2007) presents an example of Bradley-Terry models for the analysis of men's tennis for the 2004–2005 season. Using the resulting scale, it is possible to calculate the odds of victory for players who have not met in actual competition. While specialized software has been developed, it is straightforward to use the standard GLM packages to analyze data and perform a Bradley-Terry analysis. To use GLMs, choose binomial data with a logistic or probit link function. Next, the scale values are the quantities to be estimated. For each comparison, put a +1 in the column of the winning item and a -1 in the column for the losing item, with 0's elsewhere. The responses are all encoded as a 1. This formulation leads to the same results achieved by Firth (2005) and by Agresti (2007).

The method can be used on a wide variety of topics from courses of action, tone quality, technology impact, human similarity, etc., to summarize the paired comparisons in a scale. This scale will indicate the relative differences of the items and can be used to make informed decisions.

4. Tournament Method

Reviewing the red (rightmost) and green (leftmost) curves in figure 1, it is easy to imagine that comparisons will almost always favor red. Asking comparisons between red and green will not produce any mixed results unless the contest is repeated many times. For example, if asked to compare red and green five times, the result would likely be five preferences for red; there is no basis to know if the distance between the two is 3 or 10. The use of this comparison only minimally improves the overall state of knowledge. In contrast, comparisons between the blue and cyan will produce mixed results, and these probabilities form the basis for the estimation of the separation between the two. Silverstein and Farell (2001) discuss these issues and recommend that closer samples are compared more often than distant samples. This increases the value per sample over that of an exhaustive comparison.

A priori, the proximity of items is not known; if it is, then the previous suggestion can be implemented according to the prior knowledge. The Neyer method (Neyer, 1994) is used in plate penetration experiments to select the speed of impact. This method is adaptive rather than a planned experiment. Using the Neyer method, an experimenter estimates the V50 speed and the variance as an indicator of the width of the zone of mixed results. Also, as knowledge becomes available from trials, it is possible to use this information to select the experimental values to maximize information gain. In this case, there would be an adaptive experimental design. Suzuki et al. (2010) recommend a tournament system to accomplish this goal for paired comparisons. Each comparison is treated as a contest, and the goal is to find a winner. A tournament can be thought of as a method to evaluate the skill level of the contestants, typically in as few rounds as possible.

A round-robin tournament would be a full set of paired comparisons, i.e., all possible contests are played. For large tournaments, these are not used. However, for elite tournaments or invitational tournaments, this method is used. Sometimes more than one game is played between a pair of opponents. Knock-out tournaments provide a quick way to select an overall winner. For paired comparisons, the Swiss tournament system is the preferred method. The distinguishing features of a Swiss tournament are that no player is ever eliminated and, in each round, a player plays an opponent with the same number of wins. In a Swiss tournament, each round will consist of n/2 comparisons. To get a single winner, the number of rounds needed, R, can be determined from the formula $2^R = n$, where the number of rounds is the formula value rounded up to the next integer. Going back to figure 1, the effect of a Swiss tournament is to increase the comparisons between items of similar scale value. Consider 16 items for a pair-wise comparison experiment, if every comparison is made, then $\frac{16*15}{2} = 120$ comparisons will be made; however, if a Swiss system is used, R=4 and only $\frac{16}{2} * 4 = 32$ comparisons are required to obtain a winner. Note that the Swiss tournament can be extended to additional rounds to obtain more detailed information. For the example of n items, a condition on the number of rounds could be $R\frac{n}{2} \leq \frac{n(n-1)}{2}$ or $R \leq n-1$. If R>n-1, then the tournament system will require more comparisons than the full test. For a 16-item, paired-comparison experiment, a Swiss system using 15 rounds would include the same number of comparisons as the full test and provide more useful information.

5. Two-Item Discussion

To exemplify some of the issues associated with interval scale estimation, the estimation of the distance between only two items will be examined. Consider figure 1 and compare the blue and cyan ability or preference distributions. The distance between the two is 0.5, and the variance of each distribution is equal to one. The probability associated with a z-score of 0.5 is 0.6915; this is the quantity to be estimated based on binomial data from a number of trials using a probit link function. In this section, some of the issues associated with the accuracy and resolution of the method are discussed. For a difference in skill level of z = 0.5 between two items (teams, preferences, etc.), an attempt is made to demonstrate the issues associated with accuracy. First, the maximum accuracy of the probability of winning that can be achieved is determined by the number of trials. If only five trials are used the resolution of the estimate is 0.2, and the resolution in terms of the number of trials or comparisons, n, is 1/n. The number of unique probit values or distances is n+1.

For two teams separated by a probit distance of 0.5, the probability of the stronger team winning is 0.6915. Assuming that 10 comparisons are made, the binomial resolution is 0.1. There will be

11 possible outcomes for the probit function. The probit function will give the z-value that is associated with the estimated probability of winning. This z-value is the estimated distance between the two items. Figure 2 shows the probability density function (PDF) for this situation. An inspection of figure 2 reveals that while the true winning percentage is 0.6915, the closest possible estimated value is 0.7. Also, this will only occur with a probability of 0.2664.

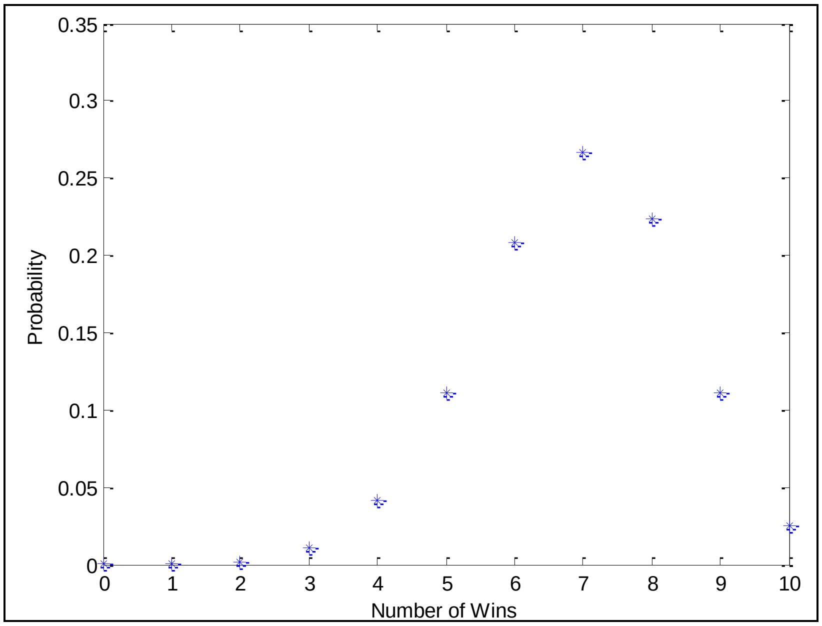

Figure 2. Binomial PDF for p(win) = 0.6915 and n = 10 trials.

Table 1 presents the precise values from figure 2. Five wins and five losses will result in an estimated difference of zero units between the two items. From table 1 and based on the data, the ordinal relationship (stronger <= weaker) will be estimated as true about 16.5% (the sum of the percentages of five or fewer wins) of the time. The third row of table 1 gives the z-values associated with the binomial result or the observed winning percentages. Each z-value is the estimated scale difference between the two items. The closest estimate to the true distance of 0.5 occurs when seven wins occur. Many packages include inverse probability functions that give the z-score associated with a given probability. The values in the third row are the only z-values or scale differences that can be estimated for this data based on the number of wins. A GLM package will not return an estimate of Inf or –Inf; typically, a relatively large magnitude is used in place of infinity. The estimate of the winning percentage will be close to the true value only 26.64% of the time. The weaker alternative will be estimated to be equal or more desirable 16.5% of the time. The analysis of tables similar to table 1 before running an experiment can be helpful in the determination the number of trials and the possible conclusions. A simulation of 10 binomial trials with p = 0.6915 was run 100 times. The estimated z-scores or probit values are displayed in figure 3. There were two cases of all wins in the 100 simulations; the estimated z-score for this situation, as given by MATLAB, was 16.295 (rather than infinity). These values are not included in figure 3 because they make other values difficult to distinguish visually.

Table 1. Key values related to figure 2.

Wins	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
P(wins)	0.0000	0.0002	0.0018	0.0106	0.0414	0.1113	0.2080	.2664	0.2239	0.1115	.0250
z(wins/n)	–Inf	–1.28	–0.84	–0.52	–0.25	0	0.25	0.52	0.84	1.28	Inf

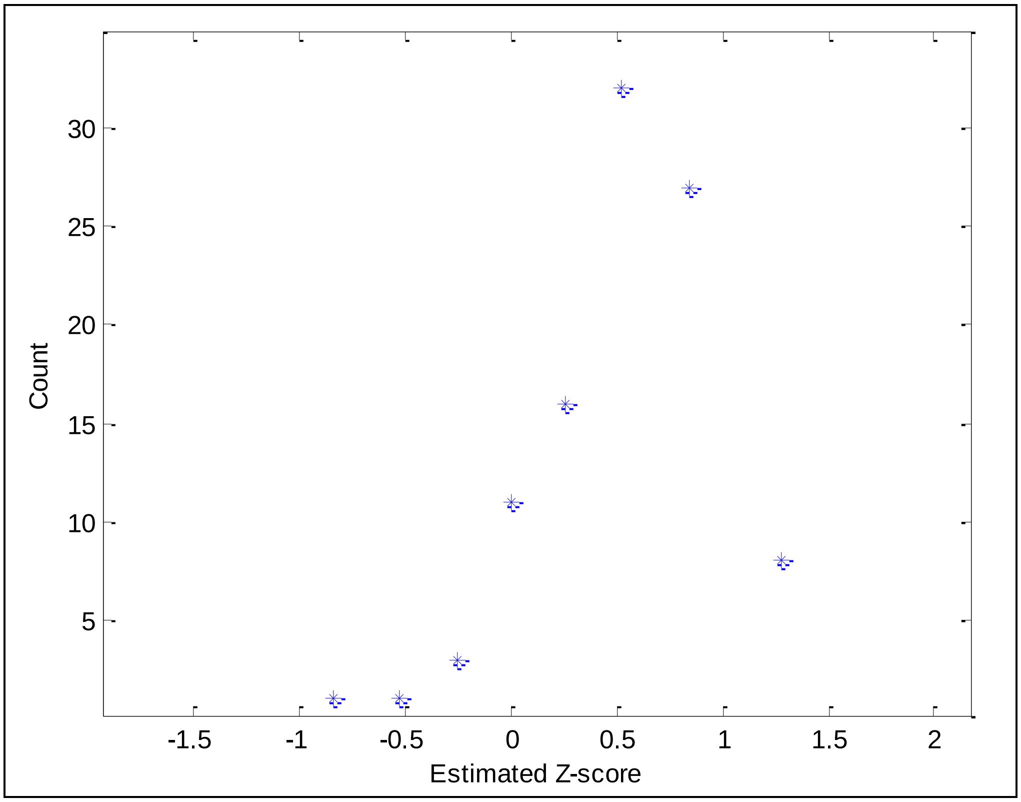

Figure 3. Simulation results for 100 replications of 10 trials.

The estimated values of the z-scores were {–0.8420 –0.5240 –0.2530 0 0.2530 0.5240 0.8420 1.2820 16.295}. The two values associated with 0 and 1 win were not realized in the simulation. Highly accurate estimates are difficult to obtain with binomial data unless the number of trials is large.

6. Comparisons of More Than Two Items

For comparisons of more than two items, the ideas presented in the previous sections are pertinent. The resolution is affected by the number of comparisons between pairs of contestants, along with the amalgamation of these contests. GLM estimation results in an overall ability scale. Simulations can be used to determine the scale accuracy achieved by a given experiment. Using prior knowledge, an investigator can assess the information gain due to different

experimental methods for paired comparisons. A problem that can arise is that the binomial trials can partition items into two sets where all the items in one set lose to all the items of the other set. This will cause large gaps in the estimated scale values between the two sets. In this situation, the distance between the two sets cannot be determined; but within each set, the estimates of scale are optimal. If this data partitioning occurs during an experiment and the goal is to form an interval scale, the experiment must allow for more comparisons between the two partitioned sets in an attempt to establish a zone of mixed results between the two sets. Another option would be to add items for comparison that seem to be in the gap between the two partitions. If possible, when designing the experiment, the investigator should attempt to select a set of items that does not contain large gaps in contest ability.

Several simulations were developed to investigate different approaches to scale development. The first simulation makes all possible comparisons; this can be considered the default position. A second simulation implements a Swiss tournament system for a specified number of rounds. The third simulation investigates the tournament system being used across similar situations. For example, the third simulation could simulate the performance of n evaluators looking at m separate systems, or comparisons of n system generated evaluations for each of m different videos.

For each of the following studies, simulations were developed using the GLM package from MATLAB. For each binomial response, the data was used with a probit link function. The results were the estimates of 16 scale values. The estimation procedure actually estimates 15 differences, so the scales can be shifted to an arbitrary starting value.

7. Study 1: Segmentation for Default and Tournament Approach

A study was performed to investigate the set partitioning problem. For this study, it was assumed that 16 items of ability increasing by 0.3 were compared. Segmentation of the estimates was investigated for the case of all possible comparisons and the tournament system with 15 rounds (the number of comparisons is the same for each design). One-hundred replications were performed for each case. For the all-possible comparisons method, 25 of 100 cases did not result in segmentation. Using the tournament method, 47 of 100 cases were not segmented. The tournament method resulted in fewer cases of segmentation. Segmentation is less likely to occur when there are more comparisons of similar ability.

8. Study 2: Scale Accuracy for Default and Tournament Approach Given No Segmentation

The next study looked at a comparison of the errors given there is no segmentation. For the tournament method (15 rounds) and the all-possible cases, 100 estimates were collected for cases of no segmentation. The sum of the square of the error between the true value and the estimated scale value was used as a measure of performance. For these cases, the squared error was accumulated over all simulations. For all possible cases, the squared error for 16 items and 100 conditional replications was 7674.7 and 1750.4 for the tournament method. This again showed the advantage of using the tournament method for paired comparisons. Scale error is reduced when there are more comparisons of items of similar ability. The tournament approach will attain accuracy goals with fewer comparisons.

9. Study 3: Scale Development Across Tasks

The final study investigated at the development of a scale for the evaluation of 16 systems that completed 480 separate tasks. For this simulation, it was assumed that the strengths of the systems varied by 0.3. First, a simulation was run for the complete set of system comparisons for each task. This was compared to a simulation of a tournament adaptive design being used for each task for each of the 480 tasks. A single scale resulted from each simulation. For the complete design, the squared residual error was 0.0175; for the tournament approach with 15 rounds, the squared residual error was 0.0104, a substantial decrease. When a tournament of four rounds was run for each task, the squared residual error was 0.0162, a slight decrease from the full test with 4/15 of the comparisons. This is slightly better performance using 27% of the paired comparisons, a significant reduction in the data acquisition effort. The advantage of using an adaptive design based on the tournament method is obvious from this study.

10. Conclusions

This report discussed some of the issues associated with the development of interval scales from paired comparisons. Each comparison can be thought of as a contest where one item wins based on a subjective or objective criterion. After scales are formulated, it is possible to make statements about the probability of each item exceeding another according to the criterion. For Turing tests, a scale developed using evaluations made by machines and people as items can be evaluated through paired comparisons on a human-like criteria. The resulting scale of human

and machine values could then be the basis of a Turing test. For decision-makers, scales can provide crucial quantitative information.

Simulations provide a useful tool for evaluating the effectiveness of different data acquisition plans. The Swiss Tournament method provides a good adaptive method for presenting paired comparisons. The fact that adaptive rather than preset experimental designs can decrease the data needed for a given level of accuracy argues for the use of adaptive designs. While no evidence is presented to indicate the tournament approach is the best possible method, the method quickly separates items into groups of similar ability and was more efficient than the exhaustive method. The development of interval scales allows a decision-maker or researcher to take a quantitative approach to subjective information.

11. References

Agresti, A. *An Introduction to Categorical Data Analysis*; 2nd ed.; Wiley & Sons: New York, NY, 2007.

Dobson, A. J.; Barnett, A. G. *An Introduction to Generalized Linear Models*; 3rd ed.; CRC Press: Boca Raton, FL, 2008.

Firth, D. Bradley-Terry Models in R. *Journal of Statistical Software* **2005**, *12* (1).

Mosteller, F. Remarks on the Method of Paired Comparisons: The Least Squares Solution Assuming Equal Standard Deviations and Equal Correlations. *Psychometrika* **1951**, 16.

Nelder, J.; Wedderburn, R. Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)* **1972**, *135* (3), 370–384.

Neyer, B. T. A D-Optimality Based Sensitivity Test. *Technometrics* **1994**, *36*, 61–70.

Silverstein, D. A.; Farrell, J. E. An Efficient Method for Paired Comparison. Hewlett Packard Laboratories. *Journal of Electron Imaging* **2001**, *10*, 394.

Suzuki, T.; Yasui, S.; Ojima, Y. Evaluating Adaptive Paired Comparison Experiments. *Frontiers in Statistical Quality Control 9;* 1st ed.; 2010; pp 341–350.

NO. OF COPIES	ORGANIZATION
1 (PDF only)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FORT BELVOIR VA 22060-6218
1	DIRECTOR US ARMY RESEARCH LAB IMNE ALC HRR 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL 2800 POWDER MILL RD ADELPHI MD 20783-1197

NO. OF COPIES	ORGANIZATION
1	DEPT OF MECHANICAL ENGRG DREXEL UNIV B C CHANG 3141 CHESTNUT ST PHILADELPHIA PA 19104
2	DARPA/I20 J FRITZGERALD P MICHELUCCI J DONLON 3701 N FAIRFAX DR ARLINGTON VA 22203
42	ORSA CORP J OLAH G THOMPSON 1003 OLD PHILADELPHIA RD ABERDEEN MD 21001
1	US ARMY RESEARCH LAB RDRL CI D DENT BLDG 205 RM 3B008 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CIS B FORNOFF BLDG 205 RM 3A014 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII A C PIERCE BLDG 202 RM 4G058 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII A S YOUNG BLDG 202 RM 4G052 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII A P DAVID BLDG 202 RM 4G062 2800 POWDER MILL RD ADELPHI MD 20783-1197

NO. OF COPIES	ORGANIZATION
1	US ARMY RESEARCH LAB RDRL CII A N FUNG BLDG 203 RM 4G054 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII B L SADLER BLDG 202 RM 4F026 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII B L TOKARCIK BLDG 202 RM 4G022 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII T C VOSS BLDG 202 RM 4G040 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CII T V HOLLAND BLDG 202 RM 4G032 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CIN S C ARNOLD BLDG 204 RM 2C116A 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	US ARMY RESEARCH LAB RDRL CIN T R HARDY BLDG 205 RM 4E008B 2800 POWDER MILL RD ADELPHI MD 20783-1197

NO. OF COPIES	ORGANIZATION
1	US ARMY RESEARCH LAB RDRL CIN T B RIVERA BLDG 205 RM 4E025B 2800 POWDER MILL RD ADELPHI MD 20783-1197
	ABERDEEN PROVING GROUND
1	DIR US ARMY EVAL CTR HQ TEAE SV PA PA THOMPSON 2ND FL 2202 ABERDEEN BLVD APG MD 21005-5001
25	DIR USARL RDRL CIH R NAMBURU RDRL CIH M D WILSON RDRL CIH N C ADAMS RDRL CII B BROOME RDRL CII C B BODT J DUMER A NEIDERER RDRL CIN D C ELLIS L MARVEL R RESCHLY RDRL HRS C E HAAS RDRL SLB D J COLLINS L MOSS RDRL SLB W P GILLICH RDRL WML A M ARTHUR B FLANDERS B OBERLE R PEARSON A THOMPSON (2 CPS) D WEBB RDRL WML F T HARKINS M ILG R MCGEE RDRL WML H T BROWN